UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/896,199	06/29/2001	Eric Cohen-solal	US010324	7568

24737 7590 01/11/2008

PHILIPS INTELLECTUAL PROPERTY & STANDARDS
P.O. BOX 3001
BRIARCLIFF MANOR, NY 10510

EXAMINER	
CHANG, KENT WU	
ART UNIT	**PAPER NUMBER**
2629	

MAIL DATE	DELIVERY MODE
01/11/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Application Number: 09/896,199
Filing Date: June 29, 2001
Appellant(s): COHEN-SOLAL, ERIC

Larry Liberchuk
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 1/19/07 appealing from the Office action mailed 8/23/06.

(1) Real Party in Interest

A statement identifying by name the real party in interest is contained in the brief.

(2) Related Appeals and Interferences

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) Status of Claims

The statement of the status of claims contained in the brief is correct.

(4) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Claimed Subject Matter

The summary of claimed subject matter contained in the brief is correct.

(6) Grounds of Rejection to be Reviewed on Appeal

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

(7) Claims Appendix

The copy of the appealed claims contained in the Appendix to the brief is correct.

(8) Evidence Relied Upon

5,999,214	INAGAKI	12-1999
6,154,723	COX	11-2000

Pavlovic et al., "Integration of audio/visual information for use in human-computer intelligent interaction", Image Processing, 1997 Proceedings IEEE, pages 121-124.

(9) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 1-18 and 20-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Inagaki (US Patent No. 5,999,214) in view of Pavlovic et al ("Integration of audio/visual information for use in human-computer intelligent interaction", Image processing, 1997 Proceedings IEEE, pages 121-124) and Cox et al (US Patent No. 6,154,723).

With regard to claim 1. Inagaki teaches a video display device comprising: a display configured to display a primary image and a picture-in-picture image (PIP) overlaying the primary image (Figure 11, items 13 and 17); and a processor operatively coupled to the display and configured to receive a first video data stream for the primary image, to receive a second video data stream for the PIP (Figure 11, items 22 and 16). Inagaki does not teach, "to recognize an audio

command related to a PIP display characteristic, the processor, upon recognizing the audio command, activates an image acquisition component that is configured to recognize a user hand gesture related to manipulating the PIP display characteristic, the processor manipulates the PIP display characteristic according to the audio command and the hand gesture". Inagaki's apparatus instead detects and responds to any of the many sounds or "audio indications" in the form of a unique voices of a specific speaking attendees with the same command which is to move the camera and highlight the PIP of the speaking attendee and does not depend on "related gesture from a user" (figure 11 "VOICE DIRECTION DETECTION UNIT", column 3, lines 31-33, column 10, lines 16-25).

However, Pavlovic demonstrates the concept of a system utilizing a combination of "audio commands" and a "related gesture" from a user as a means of controlling a graphical object on display, which is analysis to where Inagaki controlled a specific graphical object such as a PIP on a display (see Pavlovic page 123 3. Experimental Results section).

Therefore, it would have been obvious for one ordinary skill in the art at the time of the invention to use a "received audio command and related gesture from a user", as taught by Pavlovic in the apparatus of Inagaki, because of the motivation directly provided by Pavlovic: "Psychological studies, for example, show that people prefer to use hand gestures in combination with speech in a virtual environment, since they allow the user to interact without special training or special apparatus". Pavlovic further teaches that "words or gestures alone can

be used", therefore, it would have been obvious for one ordinary skill in the art at the time of the invention to use words and gestures alternatively, or simultaneously, to control the data inputting since it merely depends on the user's preference and the type of the application being used. Any levels of integration of the voice commands and gesture commands would perform equally well in providing input to the computer. Furthermore, it would have been obvious matter of design choice to choose whether to enter a voice command first, then a gesture command, or in opposite order, since it merely depends on the function being performed and the assignments of the commands. For example, in a conventional system wherein authentication is done by voice verification, a voice command from a user is obviously needed first in order to gain access to the system. In a system wherein movement of the cursor is controlled by gesture commands and selection of a menu item is input by voice commands, then whether a voice command or a gesture command is needed first would depend on the current position of the cursor: gesture commands first if the user needs to move the cursor, but voice commands first if the user wants to select the current highlighted menu item (this reads on the limitation of "the processor is configured to receive the related gesture from the user in response to the receive audio command"). As evidence of inputting a voice command before a gesture command, Cox teaches a data inputting system for a computer using voice commands and gesture commands, wherein some voice commands

trigger input from gesture commands (see column 3 lines 6-10, column 5 lines 10-19, 51-68, note that the gesture input is invoked by the voice command).

Consider claim 2. Inagaki as modified teaches the method for inputting data to a video display device having PIP windows. Therefore, it would have been obvious for one ordinary skill in the art at the time of the invention to use the data for controlling any parameter changes including size adjustment of the PIP window so as to enable simple and precise data inputting for controlling the size adjustment of the PIP window.

With regard to claim 3, Inagaki as modified teaches the video display device of claim 1, comprising a microphone for receiving the audio command from the user (See Inagaki figure 11).

With regard to claim 4, Inagaki as modified teaches the video display device of claim 1 wherein the processor is configured to analyze audio information received from the user to identify when a PIP related audio indication is intended by the user (See Inagaki figure 8a and 8b).

With regard to claim 5, Inagaki as modified teaches the video display device of claim 1, wherein the processor is configured to analyze image information received from the user after the audio command is received to identify the change in the PIP display characteristic that is expressed by the received gesture (See Inagaki figure 8a and 8b and Pavlovic et al figures 6-8 and especially the Pavlovic figure 5 "HIGH LEVEL FEATURE INTEGRATION" where it was obvious the pre analyze step is to simultaneously receive the video and audio data using

the camera and the microphone, where it is then split into a parallel visual and audio estimator/classifier module which is followed by a second stage which contains a feature integration/combination module where the combination module computes the likelihood of the pairs of gesture and verbal words. This claim language is very broad here because Pavlovic clearly receives both the audio and video before he analyzes the video or audio data, this is just the logical progression claimed).

With regard to claim 6, Inagaki as modified teaches the video display device of claim 5, wherein the image information is contained in a sequence of images and wherein the processor is configured to analyze the sequence of images to determine the received gesture (since a gesture can be a motion which would require a sequence of images to detect this feature is obvious to the system of Inagaki and Pavlovic also see Pavlovic section 21). With regard to claim 7, the combination of Inagaki and Pavlovic teaches the video display device of claim 1, wherein the image information is contained in a sequence of images and wherein the processor is configured to determine the received gesture by analyzing the sequence of images and determining a trajectory of a hand of the user (since a gesture can be a motion which would require a sequence of images to detect this feature is obvious to the system of Inagaki and Pavlovic and is merely viewed as directed towards an obvious intended use of which the combination of which it is capable also see Pavlovic section 2.1).

With regard to claim 8, Inagaki as modified teaches the video display device of claim 1, wherein the processor is configured to determine the received gesture by analyzing an image of the user and determining a posture of a hand of the user (since a gesture can be a posture of a hand this feature is obvious to the system of Inagaki and Pavlovic and is merely viewed as directed towards an obvious intended use of which the combination of which it is capable also see Pavlovic section 2.1).

With regard to claim 9, Inagaki as modified suggest the video display device of claim 1, wherein the video display device is a television (since Pavlovic shows a projection screen in figure 6 and since it is also well-known in the prior ad that televisions use projection screens one would be motivated to have a projection screen with a dual use such as conference and watching the game and is merely viewed as directed towards an obvious intended use of which the combination of which it is capable) .

With regard to claim 10, Inagaki as modified teaches the video display device of claim 1, wherein the image is a sequence of images of the user containing the user gesture, the video display device comprising a camera for acquiring the sequence of images of the user (see Inagaki figure 11, item 2).

With regard to claims 11-14, most of the limitations was already shown above with regards to apparatus claims 1-10 to be obvious and therefore the method claims 11-14 which corresponds to the apparatus were also obvious and in addition the applicant is now specifically claiming', "determining whether the

received audio command is one of a plurality of expected audio command; analyzing a gesture of the user if the received audio command is one of the plurality of expected audio indications" (SEE Pavlovic figure 7 where he illustrates a plurality of "expected audio indications" SPEECH , and a plurality of "expected gestures" GESTURE. Now look at Pavlovic figure 5 where he illustrates in the audio estimator/ classifier module receiving and "determining whether the received audio command is one of a plurality of expected audio commands" and where also he illustrates in the video estimator/classifier module receiving and "determining whether the received gesture is one of a plurality of expected gestures". It is an obvious practice that if either data collection process produces an error because the audio command or gesture used is not from the expected sets illustrated in figure 7 that the next step of "analyzing a gesture of the user if the received audio indication is one of the plurality of expected audio" in the Feature Integrator will not happen. This is because it is an obvious practice when an artificial intelligent or smart device as illustrated by the combination of Inagaki/Pavlovic can not comprehend the data within a reasonable range of certainly or as stated by Pavlovic "computes the likelihood" that it simply errors out in the flow chart and does nothing but waits for further inputs).

With regard to claims 15-18 the combination of Inagaki and Pavlovic was shown above to read on most of these limitation in claims 1-14 in addition to summarize a feature directed towards a program stored implementing this process is

inherent to the automatic computer system taught by the combination of Inagaki and Pavlovic.

With regard to claim 20, Inagaki as modified was shown above to read on these limitations in claims 1-18 (See Pavlovic figure 5 and specifically the rejection of 11 above).

With regard to claim 21, see the rejection above, note that the device of Inagaki as modified is a computer performing data inputting functions, and therefore includes the program segments for performing each of the functions.

With regard to claims 22-24, Inagaki uses a camera for image acquisition.

(10) Response to Argument

In response to appellant's arguments against the references individually, one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.*, 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986).

In response to appellant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596

(Fed. Cir. 1988)and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, Cox teaches that the voice and gesture based interaction is a highly efficient virtual direction method permitting intuitive operation by an operator in the computer input system (see column 2 lines 59-62).

The remainder of the pertinent topics for argument are present in the appropriate rejections above.

(11) Related Proceeding(s) Appendix

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

Kent Chang

Kent Chang

Conferees:

Sumati Lefkowitz

SUMATI LEFKOWITZ
SUPERVISORY PATENT EXAMINER

Amr Awad

AMR A. AWAD
SUPERVISORY PATENT EXAMINER